基于改进的 TLD 目标跟踪算法 *

胡 欣，高佳丽

(长安大学 电子与控制工程学院，西安 710064)

摘 要：针对传统跟踪—学习—检测(tracking-learning-detecting，TLD）目标跟踪算法由于检测模块扫描大量子窗口而导致检测时间过长，并且在跟踪过程中当目标发生严重遮挡、形变时，TLD 算法会出现跟踪失败的问题进行了研究，提出改进 TLD 目标跟踪算法。改进算法在检测模块前加入 ViBe 模型预估前景目标，极大地缩小了检测区域。追踪模块用 SIFT 特征匹配算法来代替原算法中的光流法，准确跟踪目标避免发生跟踪漂移，减少了计算的复杂度，提高了算法适应环境的能力。实验表明，改进后的 TLD 算法运行速度得到提升，并且当目标出现严重遮挡、光照强度剧烈变化时的跟踪精度也得到了很好的改善。

关键词：TLD 算法；ViBe 算法；SIFT 特征匹配算法；跟踪漂移

中图分类号：TP391.4 doi: 10.3969/j.issn.1001-3695.2018.05.0353

Target tracking algorithm based on improved TLD

Hu Xin, Gao Jiali

(*School of Electronics & Control Engineering*, *Chang'an University*, *Xi'an* 710064, *China*)

Abstract: Aiming at the problems of that the detecting module scans a large number of sub windows, which result the detection time is too long , and when the target has severious occlusion and deformation during the tracking process, the traditional tracking learning detection (TLD) target tracking algorithm will fail to track, so this paper propose the improved TLD target tracking algorithm. Before the detection module, it added the ViBe model to estimate the foreground target, which greatly reduces the detection area. The tracking module uses the Sift feature matching algorithm to replace the optical flow method in the original algorithm, accurately tracking the target to avoid the tracking drift, reducing the complexity of the calculation and improving the ability of the algorithm to adapt to the environment. The experiment results show that the improved TLD algorithm can improve the running speed, and the tracking accuracy can also be improved when the target is seriously occluded and the light intensity changes dramatically.

Key words: TLD algorithm; ViBe algorithm; SIFT feature matching algorithm; tracking drift

0 引言

计算机视觉广泛应用于人们的日常生活中，在计算机视觉研究领域，目标跟踪技术一直都是热点课题，在实际应用中，如何根据视频连续帧中的目标状态准确的估计出目标的位置、轨迹已成为研究数字图像处理领域的热点问题。 目前，目标跟踪技术已广泛应用于自动视频监控、汽车导航、智能机器人导航与定位、人机交互、智能交通系统、无人机目标追踪以及视频检索等领域[1]。现如今，经过很多学者的研究，视频目标跟踪技术已经相对成熟，并且取得了更大的进展，国内外的学者对目标跟踪技术的研究也越来越重视。

目前传统的跟踪算法可以分为两类，一类是基于跟踪器的目标跟踪算法，另一类是基于检测器的目标跟踪算法[2]。基于跟踪器的目标跟踪算法首先假设目标在视野中一直处于可见状态，如果目标消失或者发生遮挡，视频目标跟踪就会失败[3]，该方法具有很大的局限性，被认为是短期的跟踪器，不可以用于长时间的视频目标跟踪。基于检测器的目标跟踪算法，该方法对视频中的目标跟踪是通过滑动窗口来扫描每一帧图像，但是该算法的缺点是要提前挑选大量的被样本用来检测和学习，因此此类方法也不能很好的应用于长期目标检测[4]。在长期的目标跟踪过程中使用单纯的目标跟踪或者目标检测技术，都很难得到理想的效果。因此，考虑将目标追踪算法和目标检测算法相结合，当待追踪目标发生遮挡或者目标消失导致系统跟踪失败的时候，就可以利用检测器检测到出现的目标[5]，然后用追踪器重新进行跟踪，从而就有学者提出了 TLD 目标跟踪算法，该算法结合了传统的目标检测技术和目标跟踪技术，解决了被

收稿日期：2018-05-06；**修回日期：**2018-06-21 **基金项目：**国家自然科学基金青年基金资助项目（61701044）

作者简介：胡欣（1975-）女，副教授，博士，主要研究方向为数字图像处理（791930739@qq.com）；高佳丽（1994-），女，陕西咸阳人，硕士研究生，主要研究方向为数字图像处理.

跟踪目标在跟踪过程中发生的形变、光照、以及部分遮挡问题，并且加入了一种改进的在线学习机制[6]，从而使得整个系统更加的稳定、可靠。但是对于经典 TLD 目标追踪算法，由于检测模块扫描得到的子窗口数目过大，导致检测时间过长，影响系统运行效率。并且当目标出现严重遮挡、形变、光照变化时都会出现跟踪失败的问题，提出了本文的改进方法。

1 TLD 目标追踪算法

TLD 目标追踪算法是由 Kalal 等人[7]在 2012 年提出的一种单目标长期目标跟踪算法。该算法主要由追踪模块、学习模块、检测模块、综合模块组成[8]。算法框架结构如图 1 所示。

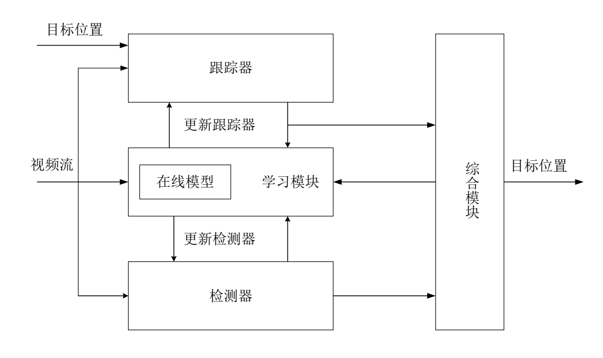


图 1 TLD 算法框架结构图

该算法的具体执行过程为：首先输入一段视频序列，在视频第一帧中手动将待追踪的目标区域定位[9]，之后并行送入检测学习追踪模块中。跟踪器通过连续两帧间的运动信息检测运动目标的状态，但是跟踪器会在目标跑出摄像头范围时跟踪失败。检测器对每一帧图像进行扫描，找到所有已经被观测学习过和目标具有类似模样的所在区域。学习模块通过对视频的在线处理提升检测器的性能，根据跟踪模块的结果对检测器的两种错误进行评估，并且根据评估结果生成训练样本对检测模块的目标模型进行更新。最后将跟踪模块和检测模块得到的追踪结果一同输入到综合模块，得到待追踪目标。

1.1 跟踪模块

为了筛选出可靠的跟踪点，跟踪器采用 LK（Lucas-Kanade）光流法改进后的中值流跟踪算法（median-flow）[10]。该算法原理图如图 2 所示，采用前向后向误差法来预测目标的位置，具有正反向连续性[11]。即从时间 t 的初始位置 $x(t)$ 开始追踪产生时间 $t+k$ 的位置 $x(t+k)$，再从位置 $x(t+k)$ 反向追踪产生时间 t 的预测位置 $x(t)$，这种前向后向轨迹预测可以得到一个位移偏差，称为前后向误差[11]。对于前后向误差值小于中值 50%的特征点，计算新的目标窗口位置，完成一次短期跟踪。

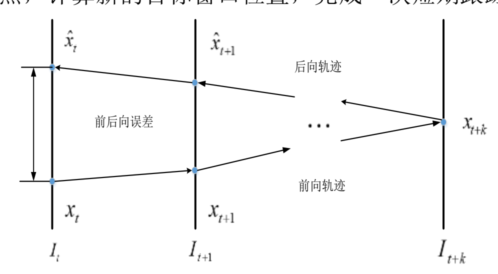


图 2 前向后向轨迹光流预测原理

1.2 学习模块

TLD 学习模块采用 P-N 半监督学习方法，通过对每一帧图像的在线处理提升检测器的性能。根据输入的检测结果，产生新的正负样本，不断更新检测器。P-N 学习的原理如图 3 所示，P-N 学习根据带标签样本监督学习训练得到初始分类器，之后利用迭代学习得到新的分类器对未标签样本数据进行分类。P-N 专家找出错误分类的样本修正训练样本集，使得下一次迭代后得到的分类器性能更好[12]。

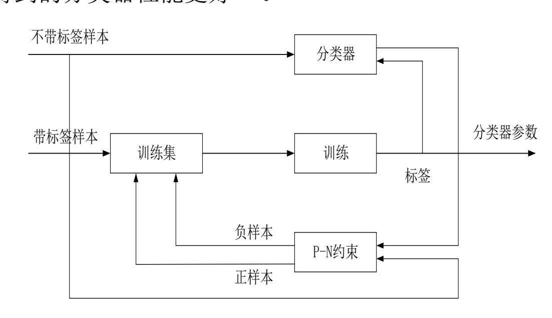


图 3 P-N 学习原理

1.3 检测模块

检测器根据目标模型，通过之前的学习和观测来定位目标在当前帧中可能出现的位置。TLD 算法中的检测器采用的是级联分类器，主要由方差滤波器、集成分类器以及最近邻分类器三部分组成[13]。通过不同尺寸的滑动窗口扫描得到的矩形框送入级联分类器。其中，方差滤波器利用积分图计算像素方差，除去方差小于被跟踪图像框方差的 50%的矩形框。剩下的矩形框送到集成分类器，平均后验概率大于 50%的矩形框则被认为是包含了目标的图像框。对于剩下的矩形框，计算其相关相似度。相关相似度大于 0.65 的即为最终的前景目标。检测器原理如图 4 所示。

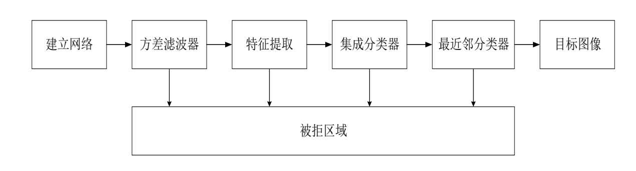


图 4 TLD 检测器原理

1.4 综合模块

TLD 综合模块将检测模块和跟踪模块得到的结果进行整合。如果两个模块都没有得到目标框，就表明当前帧中目标没有出现。否则，就将具有最大保守相似度的作为最终目标框的位置。

2 TLD 目标追踪算法的改进

2.1 检测模块的改进

经典 TLD 跟踪算法检测模块对每一帧图像进行全局扫描，产生大量子窗口，并且大多数的子窗口并不包含前景目标，这个过程使得系统的检测时间过长。其次，将整幅帧图像送入级联分类器进行检测，增加了算法的复杂度，使得整个过程的实时性变差。可以考虑采用 Vibe 算法进行前景目标检测，将含有前景目标的图像块传入方差滤波器，最终成功通过整个级联分

类器的图像块就是目标图像块，减少了检测器扫描的子窗口数目，提高了级联分类器的效率[14]。Vibe 算法是一种像素级的前景检测算法，其实时性高，检测准确。该算法的思想是：给每个像素点存储其历史像素值和邻域像素值作为样本集，对比输入新像素值和样本集判断其是否为背景点[15]。该算法主要包括三个方面：

a）初始化单帧图像中每个像素点的背景模型。选取每帧图像中像素点的 8 邻域像素值，对其随机抽样 20 次作为下一帧该像素点处的背景模型。

b）对后续的图像序列进行前景目标分割操作。计算新像素值和背景模型中采样点的距离，如果距离小于阈值的个数大于 2，就判定新的像素点为背景点，并且更新背景模型。

c）背景模型更新阶段。如果某个像素被认为是背景点，就有 1/φ 的概率对背景模型更新，更新时随机等概率的更新背景模型中的一个采样点。同时也有 1/φ 的概率更新邻域像素点的背景模型，在 8 邻域样本集中随机选择像素值进行填充更新。

2.2 追踪模块的改进

经典 TLD 算法的跟踪模块采用的是改进后的光流法，光流法使用的前提是假设亮度守恒，但在实际应用中，由于遮挡、光源、噪声的存在，亮度守恒定律一般不满足，较大的光照变化很容易造成目标丢失。其次，当目标平面旋转导致目标形变时，光流法会很难形成运动矢量场，将无法确定目标位置，导致跟踪失败。因此，跟踪模块采用 Sift 特征匹配算法来跟踪目标，该算法具有很强的匹配能力，很容易处理双帧图像，对亮度、平面旋转、目标形变具有不变性，即便是图像的拍摄角度变化很大在某种程度上也能够稳定的实现特征匹配[16]。利用该算法在实际应用中可以比较准确的追踪到目标。Sift 算法可以分解为以下四步：

a)尺度空间极值检测。二维图像 I(x,y)在不同尺度下的尺度空间可由图像 I(x,y)与高斯核函数 G(x,y,σ)卷积得到。建立图像的 DoG(Difference of Gaussian)金字塔，在 DoG 尺度空间的 26 个邻域检测极值。DOG 算子定义如下：

$$\begin{aligned} D(x,y,\sigma) &= (G(x,y,k\sigma)-G(x,y,\sigma))*I(x,y) \\ &= L(x,y,k\sigma)-L(x,y,\sigma) \end{aligned} \tag{1}$$

其中:σ 为尺度因子，L(x,y,σ)表示图像的尺度空间，k 为相邻尺度间的比例因子。

b)关键点位置及尺度确定。依据稳定程度，去除对比度较低、边缘不稳定的关键点。

c)关键点方向确定。可以通过邻域像素的梯度分布特性确定。每个点 L(x,y)梯度的模 m(x,y)和方向 θ(x,y)计算如下：

$$m(x,y)=\sqrt{[L(x+1,y)-L(x-1,y)]^2-[L(x,y+1)-L(x,y-1)]^2} \tag{2}$$

$$\theta(x,y)=\tan^{-1}\frac{L(x,y+1)-L(x,y-1)}{L(x+1,y)-L(x-1,y)} \tag{3}$$

d)关键点描述。将坐标轴旋转到特征点的主方向以保证旋转不变性。取特征点 16*16 的邻域生成该特征点的描述符，之后将 16*16 的方形区域分成 16 个 4*4 的子块，计算每个子块内 8 个方向的梯度累加值，即可形成 128 维的 SIFT 特征描述子[17]。

3 实验结果

实验以 Visual Studio 2013、OpenCV 3.0.0、MATLAB 2016a 作为开发平台，在 Windows 7 系统，处理器 Intel(R) Core^{TM} i5-6200U CPU@ 2.30 GHz 电脑上进行测试。实验共用了 Visual Tracker Benchmarks 中的 4 个测试集对 TLD 及改进 TLD 算法进行测试。表 1 对比了 TLD 算法与改进 TLD 算法的扫描窗口数量。表 2 对比了 TLD 算法与改进 TLD 算法的跟踪速度。

表 1 TLD 算法与改进 TLD 算法扫描窗口数量

测试集	TLD	改进 TLD	减少比例
David	74329	12156	6.11
Football	218965	53178	4.12
FaceOcc2	71045	8248	8.63
Jumping	99683	41763	2.38

表 2 TLD 算法与改进 TLD 算法跟踪速度

测试集	视频总帧数	TLD（帧/s）	改进 TLD（帧/s）
David	770	10.7	19.5
Football	362	12.3	18.6
FaceOcc2	812	14.3	21.8
Jumping	313	16.4	23.2

对于 David 视频集，对比原始算法和改进算法，窗口数量减少了 6 倍左右，跟踪速度由 10.7 帧/s 提升到了 19.5 帧/s；对于 Football 视频集，对比原始算法和改进算法，窗口数量减少了 4 倍左右，跟踪速度由 12.3 帧/s 提升到了 18.6 帧/s；对于 FaceOcc2 视频集，对比原始算法和改进算法，窗口数量减少了 8 倍左右，跟踪速度也由 14.3 帧/s 提升到了 21.8 帧/s；对于 Football 视频集，对比原始算法和改进算法，窗口数量减少了 4 倍左右，跟踪速度也由 16.4 帧/s 提升到了 23.2 帧/s。根据表 1 和表 2 可以看到采用了 Vibe 算法进行前景目标预估之后，大量的减少了扫描子窗口数目，并且提高了算法的运行速度，增强了整个系统的工作效率。提高系统运行速度的同时还应该考虑目标追踪的准确度，表 3 对比了 TLD 算法和改进 TLD 算法成功跟踪的目标帧数。

表 3 TLD 算法与改进算法成功跟踪目标帧数

视频库	视频总帧数	TLD	改进 TLD
David	770	628	698
Football	362	164	325
FaceOcc2	812	729	812
Jumping	313	302	313

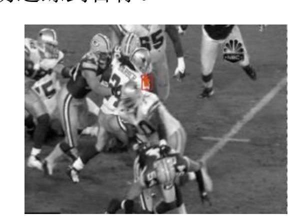
由于 SIFT 特征是图像的局部特征，其具有一定的不变性。当平面发生旋转、遮挡、尺度变化和亮度变化时，所以利用该特征匹配算法改进原始算法。对于 David 视频测试集，在追踪过程中目标发生了强烈的光照变化、以及目标姿态变化导致的平面旋转、目标被遮挡的问题，TLD 算法追踪到了 628 帧，改

进 TLD 算法追踪到 698 帧。对于 Football 测试集，在追踪过程中，目标发生了严重的形变、遮挡问题，并且目标消失重新出现在视野，TLD 算法跟踪效果较差，只成功跟踪了 164 帧，而改进 TLD 算跟踪效果较好，成功跟踪了 325 帧。对于 FaceOcc2 测试集，跟踪过程中目标发生了目标遮挡、平面旋转问题，TLD 算法追踪到 729 帧，而改进 TLD 算法成功追踪到每一帧的目标。对于 Jumping 测试集，跟踪目标出现了光照变化，摄像头抖动问题，TLD 算法成功追踪了 302 帧，改进 TLD 算法成功追踪到每一帧目标。

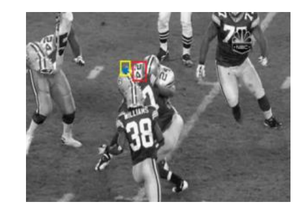
图 5 为测试集 Football 传统 TLD 算法和改进 TLD 算法追踪效果图，追踪目标是运动员队服上的号码 24。其中黄色框是 TLD 算法的追踪结果，红色框为改进 TLD 算法的追踪结果。对于此测试集视频流初始阶段，手动定位要追踪的目标。其中，（a)是在开始的第 138 帧时，两种算法都成功的追踪到了目标；（b)是在 200 帧时，运动员快速移动，导致目标出现平面变化，并且目标被严重遮挡，此时 TLD 算法跟踪目标失败，而改进 TLD 算法成功追踪到目标；（c）是第 362 帧时，追踪目标被全部遮挡且目标消失重现时，TLD 算法出现了跟踪框漂移现象，跟踪失败。而改进 TLD 算法成功追踪到目标。



(a)138 帧



(b) 200 帧



(c)362 帧

图 5　传统 TLD 算法和改进 TLD 算法追踪效果对比图

图 6 采用摄像头模式来追踪作者的嘴巴，黄色框代表 TLD 算法的追踪结果，红色框代表改进 TLD 算法追踪结果。初始化时先手动定位目标，图中，（a）是第 60 帧时中两种算法的追踪结果，可以看到 TLD 算法和改进算法都成功追踪到目标；（b）是 183 帧时作者用手遮住嘴巴，TLD 算法跟踪过程中出现跟踪框漂移到作者的脸部的现象，跟踪失败。但是改进 TLD 算法能够成功追踪到目标；（c）是当背景发生强烈光照变化时，TLD 算法跟踪失败，而改进 TLD 算法成功追踪到目标。

4　结束语

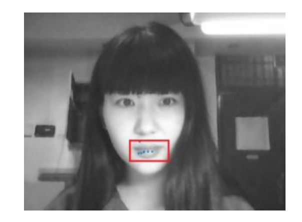
传统的 TLD 目标跟踪算法可以在先验信息较少的时候实现长时间的在线目标追踪，其跟踪速度比较快，实时性高，并且当目标消失重现后也能很好地追踪到原目标。改进后算法输入至检测模块前采用了 ViBe 算法进行前景目标预估，大量地减少了待检测的样本数量，既保证了目标的跟踪精度，还加快了检测模块的运行速度，提高了整个系统的运行效率。改进后算法在跟踪模块采用 SIFT 特征匹配算法，由于其对视角变化、仿射变换、光照变化具有一定的稳定性，所以很好地解决了原 TLD 算法的跟踪漂移问题。相对于原始的 TLD 算法，改进后算法的跟踪效果明显加强。但是本文算法性能虽然有所提升，但还存在不足之处，如在检测模块之前加入了 Vibe 前景检测算法，增加了检测模块在整个算法的比例，就会影响整个系统的协调性。而且对于学习模块，如何加强学习和提升学习精确度和灵敏度问题，都是待研究的方向。



（a）60 帧



(b) 183 帧



(c)267 帧

图 6　传统 TLD 算法和改进 TLD 算法追踪效果对比图

参考文献：

[1] 童源，费树岷，沈捷. 基于 TLD 框架的快速目标跟踪方法 [J]. 计算机应用研究, 2018, 35 (1): 317-320. (Tong Yuan, Fei Shumin, Shen Jie. Fast target tracking method based on TLD framework [J]. Application Research of Computers, 2018, 35 (1): 317-320.)

[2] 郭秋滟. 基于改进 TLD 算法的视频目标跟踪 [J]. 计算机工程与设计, 2017, 38 (9): 2551-2555. (Guo Qiuyan. Video target tracking based on improved TLD algorithm [J]. Computer Engineering and Design, 2017, 38 (9): 2551-2555.)

[3] 周鑫，钱秋朦，叶永强，等. 改进后的 TLD 视频目标跟踪方法 [J]. 中国图象图形学报, 2013, 18 (9): 1115-1123. (Zhou Xin, Qian Qiumeng, Ye Yongqiang, *et al.* Improved TLD video object tracking method [J]. Chinese Journal of Image and Graphics, 2013, 18 (9): 1115-1123.)

[4] 王铁东，任世卿. 一种改进 TLD 的目标跟踪算法 [J]. 江苏科技信息, 2018, 35 (1): 52-53, 56. (Wang Tiedong, Ren Shiqing. An improved TLD target tracking algorithm [J]. Jiangsu Science and Technology Information, 2008, 35 (1): 52-53, 56.)

[5] 周军娜，陈伟，王珂，等. 基于 TLD 的稀疏原型目标跟踪算法 [J]. 计算机工程, 2017, 43 (06): 236-240. (Zhou Junna, Chen Wei, Wang Ke, *et al.* Sparse prototype object tracking algorithm based on TLD [J]. Computer Engineering, 2017, 43 (06): 236-240.)

[6] 刘曙，狄红卫，姚曼虹. 基于自适应尺度的 TLD 目标跟踪算法 [J]. 光学技术, 2017, 43 (6): 542-546. (Liu Shu, Di Hongwei, Yao Manhong. TLD target tracking algorithm based on adaptive scale [J]. Optical Technology, 2017, 43 (6): 542-546.)

[7] Kalal Z, Mikolajczyk K, Matas J. Tracking-learning-detection [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2012, 34 (7): 1409-1422.

[8] 杨玉锋，李伟彤，许磊，等. 一种改进的 TLD 跟踪算法 [J]. 科技创新与应用, 2016 (28): 63-64. (Yang Yufeng, Li Weitong, Xu Lei, *et al.* An improved TLD tracking algorithm [J]. Technological Innovation and Application, 2016 (28): 63-64.)

[9] 曲海成，单晓晨，孟煜，等. 检测区域动态调整的 TLD 目标跟踪算法 [J]. 计算机应用, 2015, 35 (10): 2985-2989. (Qu Haicheng, Shan Xiaochen, Meng Yu, *et al.* TLD target tracking algorithm for detecting regional dynamic adjustment [J]. Computer Application, 2015, 35 (10): 2985-2989.)

[10] 王振昊. 基于 TLD 改进的人脸检测跟踪算法 [J]. 科技创新导报, 2013 (22): 50-51. (Wang Zhenhao. Improved face detection and tracking algorithm based on TLD [J]. Science and Technology Innovation Guide, 2013 (22): 50-51.)

[11] 付苗，邢藏菊. 帧间差法对 TLD 跟踪算法的改进 [J]. 电子设计工程, 2017, 25 (7): 183-186. (Fu Miao, Xing Zangju. Improvement of TLD tracking algorithm by frame difference method [J]. Electronic Design Engineering, 2017, 25 (7): 183-186.)

[12] 龚小彪. 基于 TLD 框架的目标跟踪算法研究 [D]. 成都：西南交通大学, 2014. (Gong Xiaobao. Research on target tracking algorithm based on TLD framework [D]. Chengdu: Southwest Jiaotong University, 2014.)

[13] 张丹，陈兴文，赵姝颖. 基于改进的随机森林 TLD 目标跟踪方法 [J]. 大连民族大学学报, 2016, 18 (3): 255-259. (Zhang Dan, Chen Xingwen, Zhao Shuying. Improved random forest TLD target tracking method [J]. Journal of Dalian University of Nationalities, 2016, 18 (3): 255-259)

[14] 徐勇. 基于 TLD 框架的目标跟踪算法 [D]. 南京：南京航空航天大学, 2017. (Xu Yong. Target tracking algorithm based on TLD framework [D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2017.)

[15] 刘春，翟志强. 改进的 ViBe 运动目标检测算法 [J]. 传感器与微系统, 2017, 36 (1): 123-126. (Liu Chun, Zhai Zhiqiang. Improved ViBe motion target detection algorithm [J]. Sensor and Microsystem, 2017, 36 (1): 123-126.)

[16] 冯艳. 动态背景下基于特征匹配的目标检测算法 [D]. 西安：西安电子科技大学, 2014. (Feng Yan. Target detection algorithm based on feature matching in dynamic context [D]. Xi'an: Xi'an University of Electronic Science & Technology, 2014.)

[17] 傅卫平，秦川，刘佳，等. 基于 SIFT 算法的图像目标匹配与定位 [J]. 仪器仪报, 2011, 32 (1): 163-169. (Fu Weiping, Qin Chuan, Liu Jia, *et al.* Image target matching and location based on SIFT algorithm [J]. Instrument Report, 2011, 32 (1): 163-169.)